Using Vehicular Networks to Collect Common Traffic Data

Mohammad Hadi Arbabi and Michele C. Weigle

ABSTRACT

State transportation departments are required to annually report various traffic statistics to the US government. Currently, this data is measured using older technologies that are susceptible to failure and are difficult to maintain and repair. In this paper, we propose using vehicular ad-hoc networks (VANETs) to measure this common traffic data. In addition to collecting information required for reporting purposes, VANET-based traffic monitoring can provide highly-desired metrics such as travel times, which cannot be directly measured using commonly used traffic monitoring approaches. In our system, we propose positioning a piece of infrastructure, called a task organizer (TO), near desired measurement areas. The task organizer assigns measurement tasks to equipped vehicles which then report the results back to the TO. We show that with 100% penetration rate, our methods can collect precise traffic data and that even with low penetration rates or low density traffic, we can collect high quality estimates of travel times and speeds. As travel times are difficult to obtain with currently deployed technology, we show that VANETs are well-suited for augmenting current traffic monitoring systems by providing the highly-desired travel time information.

1. INTRODUCTION

State transportation departments in the US must collect various types of data for traffic monitoring purposes. The most common of these are traffic volume, vehicle classification, traffic speed, traffic density, travel time, and vehicle miles of travel. These pieces of data are usually collected by technologies such as inductive loop detectors (ILD), video detection systems, acoustic tracking systems, and/or microwave radar sensors. In addition, wireless technologies are used in systems such as automatic vehicle identification (AVI), automatic vehicle location (AVL), and wireless location technology (WLT) [1].

ILDs, which consist of loops of wire placed in the pavement, are the most prevalent, generally have the highest accuracy, and can collect all of the fundamental traffic data except travel times. However, they are prone to failure. Maintenance, installation, and replacement can be problematic, and because of this, large portions of an ILD network may not be returning quality data at any given time. Video detection systems have issues in inclement weather (e.g., fog, rain, snow) and especially with occlusion. Occlusion is also a major problem for acoustic tracking and microwave radar systems. The ability of AVI-based systems to provide useful data is directly linked to the number of probes on the road. Therefore, these systems require the installation of significant roadside infrastructure to detect equipped vehicles. AVL systems are susceptible to the same sample size limitations as AVI systems. Also, trucking companies have been reluctant to share their AVL data with others due to concerns about losing competitive advantages in the marketplace, so data sources are not widely available. WLT systems are based on the presence of cellular phones in vehicles for monitoring traffic. But, these systems do not have the same location precision as GPS and cannot distinguish between different phones in the same vehicle.

In this paper, we propose techniques to collect common traffic data using vehicular ad-hoc networks (VANETs). We use a few pieces of roadside infrastructure to gather the information and report it to local traffic management centers. In our system, these task organizers can be deployed at various points of interest along the roadway and can be used to collect data from locations up to tens of kilometers away. We will describe how to measure and collect each of these metrics, and we analyze our proposal using simulations. Furthermore, we show the effects of penetration rate on collected data and propose an extension to the data forwarding method to improve the system for lower penetration rates and lower density traffic. Medium to high density traffic makes the usage of VANETs more feasible, and we assume for now that traffic engineers are less interested in traffic data of low density roads where VANETs may fail to collect data. Although we will study data collection in these types of conditions further, our main goal is to cover the disadvantages of traditional data collection methods and show how VANETs can be applied as a solution for medium to high density roadways.

Our goal is to develop a solution that can be used to augment currently-used systems, such as ILDs, to provide continuous high quality traffic data using a small number of roadside units. This solution must be able to collect data on dynamic length segments and at various points on the desired roads with little or no extra cost for maintenance, replacement, calibration, or infrastructure. This solution should also be able to report good estimates of traffic data in low penetration or low density traffic situations until VANET use becomes widespread.

Manuscript received ... 
The authors are with the Department of Computer Science, Old Dominion University, Norfolk, VA USA 23529-0162 (e-mail: marbab@cs.odu.edu; mweigle@cs.odu.edu).
2. RELATED WORK

There are several projects developing the use of wireless technology for traffic monitoring, such as PATH's Group-Enabled Mobility and Safety (GEMS) project [2]. GEMS is based on AVL and WLT technologies with use of Internet queries for delivering data to handheld devices. In the Mobile Millennium project [3], cell phones are the main part of the architecture. The project's concept of virtual lines is similar to our proposed virtual strips. TrafficView [4] is a scalable traffic monitoring system for inter-vehicle communication considering road conditions, but it does not consider low penetration rates or low traffic density. CarTel [5] is a distributed mobile sensor computing system that uses cell phones and cars as nodes in a dynamic sensor network. CarTel provides software to collect, process, and visualize data from sensors located on mobile devices to a central portal. Kitani et al. [6] have proposed traffic information sharing using buses on regular routes. This VANET-based technique is only useful in urban areas with good public transportation systems and only monitors those areas traveled by the transit system.

3. ARCHITECTURE

In this section, we describe the architecture of our system. The main components of the system are the roadside infrastructure and the equipment inside each vehicle. In addition, we introduce the idea of a virtual strip and describe how message forwarding is accomplished.

3.1 Requirements and Assumptions

We assume that some percentage of vehicles are equipped with a Global Positioning System (GPS) device for positioning, a detailed digital road map for route guidance, and a transceiver for communication. Dedicated Short-Range Communications (DSRC) [7] is the currently proposed standard for vehicular communications due to its low latency, making it suitable for safety applications.

We assume that there is at least one piece of roadside infrastructure, a task organizer, deployed along the road. This device is equipped with a DSRC transceiver and communicates with passing equipped vehicles. Deploying multiple task organizers along the road is an option for complex roads. The equipped vehicles and the task organizer use a common piece of application software for communication.

Through neighbor discovery (described in Section 3.4), we assume that equipped vehicles and the task organizer are aware of the positions of all equipped vehicles within the standard DSRC communication range (300 meters).

3.2 Task Organizer

A task organizer (TO) is responsible for communicating with vehicles to inform them about upcoming traffic conditions, assign measurement tasks, collect traffic data and organize the received measurements. The TO may be the property of the local Department of Transportation (DOT) and should be able to directly communicate with the local Traffic Operations Center (TOC). The goal of the TO is to provide accurate measurement information to the TOC and to disseminate timely messages from the TOC to equipped vehicles.

3.3 Vehicles

Equipped vehicles contain a GPS device, DSRC transceiver, and a detailed digital map. They are able to record their instantaneous speed, acceleration, spatial location, direction, and route (i.e., highway) number along with a timestamp. Each transmitted message from a vehicle contains a header that includes all of above information. Vehicles may receive tasks from a TO and forward the tasks to other vehicles. They can also produce new messages and forward them back to the TO. Vehicles can communicate with the TO directly or indirectly (via messages forwarded to the TO by other vehicles). Vehicles have the ability to store events that should be fired at a specific time, speed, or location. As vehicles will be running other VANET applications, we anticipate that much of the information needed to report traffic measurements can be piggybacked on the packets from these other applications or be gathered from data produced by these other applications, as is done by Robinson et al.’s Message Dispatcher [8]. We assume vehicles and TOs are able to communicate securely using the latest security techniques introduced for VANETs [9, 10, 11, 12, 13].

3.4 Neighbor Discovery, Geographical Routing, and Message Forwarding

The system can use well-known neighbor discovery, geographical routing, and message forwarding techniques (such as [14, 15]) to pass tasks and information through the VANET and to the TO. Our focus in this paper is on how to gather traffic measurement data rather than introducing a new forwarding or routing algorithm. Both vehicles and the TO perform neighbor discovery periodically, for instance each second. In neighbor discovery, each equipped vehicle broadcasts a report containing its current position, speed, direction, route, ID, and timestamp. This type of neighbor discovery may be part of a larger geographical routing algorithm.

3.5 Virtual Strips

A virtual strip is an imaginary line that crosses a road and is used to define the measurement areas. Virtual strips, or strips, can be defined geometrically as the intersection of \( w(x, y, z) = ax + by + cz = 0 \) (a plane) with the road. A vehicle can reside on only one side of a strip. If a vehicle is moving towards a strip, it is before the strip, and once it passes the strip it is considered to be after the strip. Two virtual strips can be used to create a virtual segment, or
segment. Vehicles can be inside the segment or outside the segment.

4. COLLECTING TRAFFIC DATA

This section describes how various traffic metrics can be collected using equipped vehicles and task organizers.

4.1 Traffic Volume, Speed, Classification, and Vehicle Miles Traveled

We can use the same basic procedure to gather several important traffic metrics: traffic volume, time mean speed (TMS), vehicle classification, and vehicle miles traveled (VMT).

Traffic volume refers to the number of vehicles that cross a specific point on the road in a particular amount of time. TMS is the average of the speeds of all vehicles that pass a particular point on the road in a given amount of time. The TO can compute TMS at the same time as it computes volume because it knows the total number of vehicles that have passed the desired strip and their speeds. Vehicle classification data records traffic volume with respect to the type of vehicle that passes a particular point on the road. The Federal Highway Administration has defined a set of 13 vehicle classes that are commonly used by most states [1, 16]. We assume that vehicle classification can be set either manually or automatically for vehicles that can have several classifications (e.g., a heavy truck with cabin and container). VMT is the product of the traffic volume and the length of the segment being examined. The TO can compute this metric for a specified segment that has no entrances or exits once it knows the traffic volume at the first strip of the segment.

### Pseudocode 1. Collecting traffic volume.

```
CalcVolume [speed, classification, …]
V - current vehicle
P - current vehicle position
TO - task organizer
w_j - target strip
K - task (contains TO, w_j)
E - vehicle’s event list
M - message

Receive volume task K from TO
if (K ∈ E) ignore, exit
else add K to E
while (P < w_j)
  update GPS position P
if (P > w_j)
  M = "VOL", current speed, classification, timestamp, ID
Forward M to TO
Remove K from E
```

In Pseudocode 1, we outline our algorithm to collect the traffic volume at a desired virtual strip. This method also collects the TMS and vehicle classification. (VMT can be calculated once the traffic volume is known.) The TO broadcasts a "Volume" task containing the location of the TO and the strip of interest, w_j. Each vehicle passing the TO receives the "Volume" task. Once the vehicle passes strip w_j, the task is triggered, and the vehicle sends a message back to the TO containing the "VOL" tag, its current speed, and its vehicle classification. Once several messages have been received, the TO can calculate the volume, TMS, vehicle classification, and VMT metrics.

4.1.1 Target Strip Outside Range of TO

First, we consider the scenario where the target strip is outside the communication range of the nearest TO. Through neighbor discovery, the TO knows about the vehicles that are within DSRC range and before the TO’s nearest virtual strip. The TO sends a task to these vehicles, requesting them to forward volume information back to the TO once they pass the target virtual strip. These vehicles store the task as an event that should be raised when the vehicle passes the location of the target strip. Each time the TO receives a completed volume message, it will increment the number of vehicles that have passed the target strip. Thus it can compute the traffic volume at the target strip for any duration.

4.1.2 Target Strip Within Range of TO

If the target strip is within communications range of the TO, no forwarding will be required. In this case, the TO may be able to collect volume, speed, and classification data for each lane of the road. The TO can use its own location as a reference and the location of each vehicle that passes the target strip as a target location. Therefore, it can estimate the lane in which the vehicle is traveling. There might be some GPS inaccuracy when a vehicle reports its location, but the TO can approximate this inaccuracy by assuming a particular width for each lane and comparing the vehicle's reported location to the location of vehicles traveling in the adjacent lanes as a second reference. It is the TO’s responsibility to collect the coordinates of vehicles during some period to estimate the boundaries of the left-most and right-most lanes. The received signal strength indicator (RSSI) may also be used to estimate the lane number, as it has been shown that RSSI can be used by a receiver to estimate the location of the transmitter [17].

4.1.3 Low Density or Low Penetration Rate Considerations

In situations where there are few equipped vehicles (either due to low penetration rate or low density traffic), message forwarding may fail, and the TO may count a much lower volume and density than the ground truth. There are ways that the TO can determine the difference
between low density traffic and low penetration rate by examining the speed limit of the segment, the density of the segment, and the current vehicles’ speeds. Low density traffic usually results in a higher average vehicle speed (vehicles are traveling in free-flow). Low penetration rates in mid-to-high density traffic will likely have a lower average speed due to transient periods of congestion. Road profiles and the past history of the roadway may also be a great help in determining the difference between low penetration rates and low traffic density. There are several papers that have investigated the impact of low density traffic (resulting in a sparse network) in VANETs [14, 15, 16, 18, 19]. We suggest measuring traffic volume at some point inside the communication range of a TO to avoid the need for forwarding and to mitigate the effects of low traffic density.

4.2 Traffic Density

Traffic density refers to the number of vehicles on a section of road. Current methods that collect this type of data, like aerial photography, are not cost effective. Most often, detector occupancy [1], the percentage of time that the detector (such as an ILD) is active due to the presence of a vehicle, is used as a surrogate measure for traffic density. We show how VANETs can be used to collect instantaneous, accurate traffic density of a desired segment on the road. In the following explanation, we assume a limited deployment with a single TO in order to minimize cost. The existence of additional TOs would necessarily reduce the amount of forwarding required as the resulting density measurement could be passed to an upcoming TO rather than being passed backwards to the originating TO.

The basic algorithm is outlined in Pseudocode 2. The TO periodically selects a vehicle that is in range and sends it a "Density" task containing the location of the TO, the boundaries of the segment of interest (\(w_1\) and \(w_2\)), and a count initialized to 1. If the vehicle is inside the segment of interest, the task is triggered. The vehicle counts the number of its neighbors that are both inside its communication range \(R\) and before \(w_2\). Then, the vehicle updates the count in the task message. If the distance to \(w_2\) is greater than \(R\), the vehicle will forward the task (with the updated count) to its farthest neighbor inside \(R\) and before \(w_2\). Otherwise, the updated task message will be forwarded back to the TO. If the vehicle receiving the task from the TO is outside the segment of interest, it will initiate message forwarding to a vehicle either at the edge of its communication range or to a vehicle just inside the segment of interest.

Vehicles in charge of counting the number of vehicles can also aggregate the average speed of the vehicles they count. In this case, the same message that contains density information can contain the average speed of vehicles counted inside the segment. This average speed may be another useful metric that the TO can collect. The TO could request this kind of task periodically to monitor the current travel situation on the roadway.

<table>
<thead>
<tr>
<th>CalcDensity</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>TO</strong> - task organizer</td>
</tr>
<tr>
<td><strong>C</strong> - vehicle count inside target segment</td>
</tr>
<tr>
<td><strong>w_1</strong> - target strip [segment beginning]</td>
</tr>
<tr>
<td><strong>w_2</strong> - target strip [segment end]</td>
</tr>
<tr>
<td><strong>K</strong> - task (contains <strong>TO</strong>, <strong>w_1</strong>, <strong>w_2</strong>, (C = 1))</td>
</tr>
<tr>
<td><strong>E</strong> - event list</td>
</tr>
<tr>
<td><strong>N</strong> - neighbor list</td>
</tr>
<tr>
<td><strong>V</strong> - current vehicle</td>
</tr>
<tr>
<td><strong>P</strong> - current vehicle position</td>
</tr>
<tr>
<td><strong>d</strong>((x, y)) - distance from (x) to (y)</td>
</tr>
<tr>
<td><strong>R</strong> - communication range</td>
</tr>
</tbody>
</table>

Receive density task \(K\) from **TO**
if \((K \in E)\) ignore, exit
else add \(K\) to \(E\)
while \((P < w_1)\)
  update GPS position \(P\)
if \((P > w_1)\)
  if \((P < w_2)\) Collect \((TO, w_1, w_2, C)\)
else
  \(D = d(P, w_1)\)
  \(R = \min(R, D)\)
  \(N' \in N \cap d(V, N') \leq R\)
  \(V' \in N' \cap d_{\max}(V, V')\)
  Forward \(K\) to \(V'\)

Collect:
\(D = d(P, w_2)\)
\(N' \in N \cap d(V, N') \leq R\) and \(N' < w_2\)
\(C += \# N'\)
if \((D > R)\)
  \(V' \in N' \cap d_{\max}(V, V')\)
  Forward modified \(K\) to \(V'\)
else
  Forward \(K\) to **TO**
  Remove \(K\) from \(E\)

Pseudocode 2. Collecting density of a segment.

In our description, we have assumed a 100% penetration rate of equipped vehicles as our method can only count equipped vehicles. With a lower penetration rate, the TO can estimate density if it knows the approximate penetration rate ahead of time.

It is possible that messages cannot be forwarded to measure density in an entire segment due to low density traffic or an obstacle blocking the roadway. In such situations, if a delegate vehicle cannot forward the message farther, it will forward to the TO the density up to the last strip in which it was able to collect data. In the worst case, the TO may not receive any message for a requested density.
task. If this occurs repeatedly over some time interval, the TO should check the traffic volume to infer either low density traffic or severe congestion inside the segment.

4.3 Travel Time

Travel time is the amount of time that takes a vehicle to travel between two points on the road. It is the piece of data most understandable for the driving public and, thus, is the most desired data for traffic engineers. Unfortunately, gathering travel times has been challenging [1].

We propose a method to collect travel times, which is very similar to the method used to collect volume data. In general, if vehicles include a timestamp when they report the volume message, the travel time is the difference between the timestamps gathered from two different strips for each vehicle. Thus, the TO only needs to keep a record of vehicles that have passed two particular strips. When it is collecting volume for those strips, the TO can also compute the travel time for an individual vehicle, or the mean travel time for all reporting vehicles over a certain period of time.

An alternate method for calculating the travel time is to let each vehicle compute its travel time after it passes the two strips of the segment and send the result back to the TO.

Another metric of interest to traffic engineers is the space mean speed (SMS). SMS is based on the average speeds over an extended segment. To compute SMS, the TO needs the travel time of each vehicle that passes through a segment. Then, the TO can calculate the SMS as the length of the segment divided by the average travel time.

5. EVALUATION

We have designed several simulations to show how our proposed techniques can be used to collect common traffic data. We also use these simulations to analyze the advantages and disadvantages of using VANETs in comparison to conventional measurement technologies currently in use.

5.1 Methodology

In our simulation, developed using the ns-3 network simulator [20], vehicles are equipped with a DSRC transceiver with transmission range of $R$ ($300 \text{ m}$). Vehicles can raise events according to the task they receive or tasks they have stored. There is continuous interaction between vehicle’s transceiver, GPS device, and its controller program. The mobility of vehicles, including lane changes, is controlled by a revised version of the IDM/MOBIL car-following model [21, 22, 23]. In addition, the user can control the vehicle, causing it to slow down or stop, which can result in congestion on the roadway. The roadway is a multi-lane bi-directional road with entrances and exits. There may be several TOs next to the road, but we use only one TO in our simulations. When evaluating the accuracy of data collection, we will compare the gathered data to the actual simulation state of the road.

We examine a section of road with $10 \text{ km}$ length. The road is a bi-directional roadway with three lanes in each direction. Vehicles enter the road with a speed less than or equal to the speed of vehicles in front of them and with a fixed or variant offset from the front vehicle. Two parameters in IDM that explicitly affect the gap between vehicles are desired velocity $v_0$ and time headway $T$. We consider $v_0$ as the speed limit for the roadway. $T$ represents the minimum time spacing between two vehicles in the same lane.

Car-following models describe the lane-level gap, the headway gap between two adjacent vehicles in the same lane. From the network connectivity standpoint, however, the most relevant metric is the road-level gap, the gap from the leading vehicle to the nearest following vehicle on a multi-lane road, regardless of whether the following vehicle is on the same lane or on a different lane from the leading vehicle [18]. The volume and density of a segment of the road are dependent on both the average lane-level gap and the average road-level gap between vehicles inside the segment. These metrics also depend upon the average speed, which is a result of the parameters $v_0$ and $T$.

For all of the parameters in the IDM model, besides $v_0$ and $T$, we use the same parameters as found in the source code of Treiber’s IDM demonstration applet (specifically in constants.java) [23, 24, 25]. As with the demonstration applet, cars and trucks have slightly different parameters (e.g., trucks have lower $v_0$ and lower acceleration and deceleration rates than cars). In the experiments we present here, we used a mixture of 80% cars and 20% trucks, although we found that different mixtures produced similar results.

We examine a section of road with $10 \text{ km}$ length. The road is a bi-directional roadway with three lanes in each direction. Vehicles enter the road with a speed less than or equal to the speed of vehicles in front of them and with a fixed or variant offset from the front vehicle. Two parameters in IDM that explicitly affect the gap between vehicles are desired velocity $v_0$ and time headway $T$. We consider $v_0$ as the speed limit for the roadway. $T$ represents the minimum time spacing between two vehicles in the same lane.

Car-following models describe the lane-level gap, the headway gap between two adjacent vehicles in the same lane. From the network connectivity standpoint, however, the most relevant metric is the road-level gap, the gap from the leading vehicle to the nearest following vehicle on a multi-lane road, regardless of whether the following vehicle is on the same lane or on a different lane from the leading vehicle [18]. The volume and density of a segment of the road are dependent on both the average lane-level gap and the average road-level gap between vehicles inside the segment. These metrics also depend upon the average speed, which is a result of the parameters $v_0$ and $T$.

For all of the parameters in the IDM model, besides $v_0$ and $T$, we use the same parameters as found in the source code of Treiber’s IDM demonstration applet (specifically in constants.java) [23, 24, 25]. As with the demonstration applet, cars and trucks have slightly different parameters (e.g., trucks have lower $v_0$ and lower acceleration and deceleration rates than cars). In the experiments we present here, we used a mixture of 80% cars and 20% trucks, although we found that different mixtures produced similar results.

5.2 Measuring Volume, TMS, SMS, and Travel Time

In this set of experiments, our goal was to evaluate the effectiveness of our techniques in collecting volume, TMS,
SMS, and travel times. The TO broadcasts the desired task every second. The task requests that each vehicle send its own data back to the TO when passing strips $w_2$, $w_5$, and $w_9$, which are 1, 4 and 8 km away from the TO, respectively. Data sent by each vehicle contains its timestamp, classification, speed, and travel time for the desired segment.

To simulate different traffic densities, we generated scenarios with different average road-level gaps by varying the vehicle spacing offset, $v_0$, and $T$ parameters for IDM/MOBIL.$^1$ To measure the effectiveness of our techniques, we record the percentage of vehicles' task response messages that never reached the TO. The quality of collected and calculated traffic data by the TO are proportional to the amount of received data, which is a function of the selected routing algorithm. We aim to collect every metric generated by vehicles, similar to what ILDs do. Approximation and summary of traffic data can be a post-processing step for a TO. In this paper, we are not interested in evaluating the forwarding/routing algorithm, but rather in showing how a VANET can be used to collect the metrics.

Figure 2 shows the percentage of messages generated by vehicles passing the desired strips that were never received by the TO with different traffic densities and different speed limits ($v_0$). When the average road-level gap is well-below the communication range of 300 m, there is no problem with messages being received by the TO. With an average road-level gap of 150 m or less, almost all messages were received. Thus, the calculated volume, travel time, TMS, and SMS were identical to the values obtained by examining the simulation status (i.e., the actual values of these metrics). For an average road-level gap between 150-300 m, the unreceived message percentage increases, which means that some vehicles’ reports were not received. Once the road-level gap reached 300 m, many messages were lost since the network was essentially disconnected.

Unreported volume messages definitely impact the volume and classification data. The TMS may be affected slightly but these differences are negligible since we generate free-flow traffic with no congestion. Thus, missing some messages will not affect the calculated average speed. Likewise, since SMS and travel times are averages, missing a few messages will not affect the computed values greatly.

5.3 Response Message Delay

In addition to the accuracy of the metrics calculated by the TO, we are also interested in their timeliness. Because each vehicle timestamps its response message, the TO can calculate the average delay from when a response message was generated to when the TO received it.$^2$ Figure 3 shows the response message delay according to the distance of the vehicle from the TO. It only takes 1.6 seconds to receive messages from 9 km away. Messages reporting on data within 5 km from the TO return in less than 1 second.

5.4 Measuring Traffic Density

In this set of experiments, the goal was to evaluate the effectiveness of our technique at computing traffic density. We tasked vehicles with measuring the density between segments $w_2$-$w_5$ and $w_5$-$w_8$. The TO selects one vehicle in communication range closest to strip $w_1$ (the TO’s strip) and assigns it the density task. This action can occur periodically to allow the TO to receive multiple estimates of the traffic density. In our simulation, the TO sends a new density task message every 10 seconds. As the penetration rate was still assumed to be 100%, the percentage of

---

$^1$ We ran extensive simulations to determine the appropriate values of these parameters to produce the desired average road-level gap. For brevity, these results have been omitted.

$^2$ We assume that vehicles and the TO have synchronized clocks through GPS. If not available, the vehicles’ clocks could be synchronized with the TO upon receiving a task message.
unreceived response messages was similar to that shown in Figure 2.

Figure 4 shows the comparison between the actual density and the instantaneous reported density for each interval that the TO requested the density task. We examined two different traffic scenarios in this case. In both cases, \( v_0 = 25 \, \text{m/s} \) and \( T = 0.5 \). We varied the vehicle spacing offset to produce different traffic density levels. In the medium density case, the offset was set to 5 m to produce an average road-level gap around 25-30 m, and in the low density case, the offset was set to 500 m to produce an average road-level gap around 200 m. Note that there are two unreceived density messages due to low density traffic, one at 40 seconds and the other at 150 seconds. Nonetheless, all received messages carried the correct instantaneous traffic density.

5.5 Effect of Lower Penetration Rates

In our previous experiments, we assumed that all vehicles were equipped with a DSRC transceiver (100% penetration). They received tasks and replied to the requests as needed. We observed that all messages were generated appropriately and that if some response messages did not reach the TO it was due to network disconnectivity.

For this set of experiments, we aim to investigate the effect of penetration rate on the accuracy of our techniques. To achieve this, we keep all simulation settings the same as in the previous experiments except for the penetration rate. We turn off vehicles’ transceivers at random to generate traffic with 50%, 25%, or 5% equipped vehicle penetration, which means that 50%, 75%, or 95% percent of vehicles, respectively, are not able to communicate. We measure the percentage of messages generated by equipped vehicles that never reached the TO. This percentage is expected to be higher than with the 100% penetration rate as there will be higher network disconnectivity.

Figure 5 shows the percentage of unreceived response messages as the penetration rate changes with \( v_0 \) set to 25 m/s. All generated messages are received by the TO when the average road-level gap is less than 50 m with 50% penetration. But, we know that only 50% of vehicles are able to generate messages, so the TO only receives volume information from half of the total vehicles.

For higher road-level gaps such as 150 m, the TO does not receive 60% of generated messages due to network disconnectivity. Therefore, with 50% penetration the TO will miss 80% of the total information (0.5 + 0.6×0.5) in comparison to the actual roadway status. The amount of information the TO receives will decrease as the penetration rate decreases. In these cases, good results are only obtained when the average road-level gap is low. We investigate ways to mitigate this problem in Section 5.6. We note that we ran these same experiments with different values for \( v_0 \) (ranging from 5 m/s to 30 m/s), but the results were similar to those shown in Figure 5.

5.5 Effect of Lower Penetration Rates

In our previous experiments, we assumed that all vehicles were equipped with a DSRC transceiver (100% penetration). They received tasks and replied to the requests as needed. We observed that all messages were generated appropriately and that if some response messages did not reach the TO it was due to network disconnectivity.

For this set of experiments, we aim to investigate the effect of penetration rate on the accuracy of our techniques. To achieve this, we keep all simulation settings the same as in the previous experiments except for the penetration rate. We turn off vehicles’ transceivers at random to generate traffic with 50%, 25%, or 5% equipped vehicle penetration, which means that 50%, 75%, or 95% percent of vehicles, respectively, are not able to communicate. We measure the percentage of messages generated by equipped vehicles that never reached the TO. This percentage is expected to be higher than with the 100% penetration rate as there will be higher network disconnectivity.

Figure 5 shows the percentage of unreceived response messages as the penetration rate changes with \( v_0 \) set to 25 m/s. All generated messages are received by the TO when the average road-level gap is less than 50 m with 50% penetration. But, we know that only 50% of vehicles are able to generate messages, so the TO only receives volume information from half of the total vehicles.

For higher road-level gaps such as 150 m, the TO does not receive 60% of generated messages due to network disconnectivity. Therefore, with 50% penetration the TO will miss 80% of the total information (0.5 + 0.6×0.5) in comparison to the actual roadway status. The amount of information the TO receives will decrease as the penetration rate decreases. In these cases, good results are only obtained when the average road-level gap is low. We investigate ways to mitigate this problem in Section 5.6. We note that we ran these same experiments with different values for \( v_0 \) (ranging from 5 m/s to 30 m/s), but the results were similar to those shown in Figure 5.
Table 1. Calculated average travel time and percentage of difference for segment $w_5$.

<table>
<thead>
<tr>
<th>$v_0$ (m/s)</th>
<th>Travel Time (s) at 100%</th>
<th>Travel Time (s) at 5%</th>
<th>Difference %</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>625</td>
<td>624.7</td>
<td>0.048</td>
</tr>
<tr>
<td>10</td>
<td>309.28</td>
<td>307.6</td>
<td>0.543</td>
</tr>
<tr>
<td>15</td>
<td>203.4</td>
<td>202.1</td>
<td>0.639</td>
</tr>
<tr>
<td>20</td>
<td>158.73</td>
<td>157.9</td>
<td>0.523</td>
</tr>
<tr>
<td>25</td>
<td>126.58</td>
<td>124.6</td>
<td>1.564</td>
</tr>
<tr>
<td>30</td>
<td>106.76</td>
<td>105.2</td>
<td>1.461</td>
</tr>
</tbody>
</table>

5.6 Mitigating the Effect of Low Penetration Rates

As discussed in the previous section, lower penetration rates will impact the percentage of received information. Not only are non-equipped vehicles unable to report the desired information, but also reports from equipped vehicles could be lost during forwarding toward the TO due to the effect of penetration on network connectivity. To mitigate this loss, we modify the message forwarding method and investigate the amount of improvement we might obtain.

The modification to the forwarding algorithm consists of two parts:

- Vehicles may forward messages to equipped vehicles traveling in the opposite direction in order to route messages toward the TO.
- If a message cannot be forwarded farther toward the TO, then the equipped vehicle in the opposite direction (traveling towards the TO) will carry the message until it can deliver the message directly to the TO.

We ran simulations with settings similar to those in Section 5.5 except with the modified forwarding algorithm. The traffic flow and penetration rate in the opposite direction is the same as in the forward direction.

Figure 6 shows the results of these simulations as compared to the original forwarding algorithm. There is little that can be done when the penetration rate is as low as 5%. But, the percentage of unreceived messages is actually lower with the modified forwarding method at 25% penetration than with the original method at 50% penetration. This is because using the traffic in the opposite direction improves forwarding and reduces network disconnections. In addition, allowing an equipped vehicle in the opposite direction to carry the message directly to the TO when network is disconnected improves the TO's reception rate. The probability of finding such an equipped vehicle decreases when the penetration rate becomes lower.

In addition to using a modified forwarding algorithm, any increase to the communication range would improve performance in low penetration rate or low density scenarios. We ran a set of experiments with the communications range extended to 1000 m and found that the percentage of unreceived messages decreased with the extended range due to the lower probability of network disconnection. Figures have been omitted for brevity.

5.7 Road Exit and Entrance

We ran additional simulations with an exit at the mid-point of the road (strip $w_5$) which randomly remove $X\%$ of the vehicles in the rightmost lane from the roadway. $X$ was set to 22, 33, and 66 for three different runs. We observed that unreceived message rates remained roughly similar to previously obtained results and the average road-level gap was only increased by 1.5-4.5%. We emphasize that the quality of calculated traffic statistics and the amount of useful data collected by the TO depends on the amount received data, which is a function of road level gap, communication range, and routing algorithm.

We obtained similar results when we injected vehicles into the roadway at strip $w_5$. We used only one TO at strip $w_1$ to collect the exit and entrance data. It is possible to use TOs deployed at the exit and entrance areas to collect the traffic data of each branch of the roadway.

6. CONCLUSIONS AND FUTURE WORK

In this paper, we proposed techniques to collect common traffic data using vehicular ad-hoc networks (VANETs). We showed that a single task organizer (TO) is sufficient to collect data from up to 10 km away. Data can be collected from dynamic segments of the road and the use of virtual strips adds no extra expense. We showed the effects of penetration rate and traffic density on collected data, and we showed how bi-directional mobility (message forwarding by traffic in both directions) can enhance the proposed system. Our results showed that a TO can collect...
precise data with 100% penetration and collect good estimates of traffic data such as travel time, TMS, and SMS even in low penetration rates and with low density traffic in highways. As travel times are difficult to obtain with currently deployed technology, VANETs are well-suited for augmenting current traffic monitoring systems by providing travel time information. Although a TO can only measure the density and volume of equipped vehicles, it may be possible to approximate these metrics by having a priori knowledge about the likely penetration rate.

In future work, we plan to investigate the impact that traffic congestion and incidents have on our data gathering techniques and to investigate the impact of using additional TOs along the highway. We also plan to investigate how density and volume can be approximated in situations of low penetration rates. Further, we plan to investigate the suitability of our techniques for measuring traffic data on urban and arterial roads.

7. ACKNOWLEDGMENTS

We thank Dr. Michael Fontaine of the Virginia Transportation Research Council for helpful comments and discussions.

8. REFERENCES
